The Collective Stance in Modeling Expertise in Individuals and Organizations

Brian R. Gaines
University of Calgary

ABSTRACT: This paper is concerned with modeling the nature of expertise and its role in society in relation to research on expert systems and enterprise models. It argues for the adoption of a *collective stance* in which the human species is viewed as a single organism recursively partitioned in space and time into suborganisms that are similar to the whole. These parts include societies, organizations, groups, individuals, roles, and neurological functions. Notions of expertise arise because the organism adapts as a whole through adaptation of its interacting parts. The phenomena of expertise correspond to those leading to distribution of tasks and functional differentiation of the parts. The mechanism is one of positive feedback from parts of the organism allocating resources for action to other parts on the basis of those latter parts past performance of similar activities. Distribution and differentiation follow if performance is rewarded, and low performers of tasks, being excluded by the feedback mechanism from opportunities for performance of those tasks, seek out alternative tasks where there is less competition. The knowledge-level phenomena of expertise, such as meaning and its representation in language and overt knowledge, arise as byproducts of the communication, coordination, and modeling processes associated with the basic exchange-theoretic behavioral model. The model is linked to existing analyses of human action and knowledge in biology, psychology, sociology, and philosophy, and is used to analyze the role of information technology in supporting activities in the lifeworld.

1. INTRODUCTION

This paper continues a line of research that focuses on the modeling of expertise, its formation and the role of knowledge in expertise. Previous results have shown that positive feedback

International Journal of Expert Systems Pp. 19–49, Volume 7, Number 1, ISSN 0894-9077
Copyright © 1994 by JAI Press Inc. All rights of reproduction in any form reserved.

processes in the social management of expertise are sufficient to account for an initially homogeneous pool of learning agents to become highly differentiated in terms of their relative expertise (Gaines, 1987, 1988). These results apply to arbitrary populations of learning agents from neural complexes, through people, to organizations (Gaines, 1991). The positive feedback processes tend to concentrate expertise within individual agents, and the papers cited present expertise dissemination activities as counterbalancing processes that make that expertise more widely available. From a functional perspective, it is reasonable that an adaptive organism consisting of large numbers of short-lived, fragile, and similar agents should achieve functional differentiation through one mechanism that supports the creation of individual expertise, and also achieve independence of volatile individual components by a second mechanism that supports the dissemination of that expertise.

From an expert systems perspective, one important feature of this formulation is that it models the essence of individual expertise as being both the adaptivity to develop high performance, and the capability to deliver that high performance. An 'expert' is a dynamic adaptive system—not just a static performance system (Clancey, 1989). A second important feature is that neither the mechanisms of expertise acquisition nor those of expertise dissemination are specified by the model. A very wide variety of mechanisms are possible and sufficient to account for the phenomena of expertise. In particular, the concept of 'knowledge' is secondary to that of expertise, and the emphasis on the 'knowledge level' in much research and development relating to expert systems provides an overly narrow focus that is insufficient to account for major phenomena of human expertise. A third important feature is the essential social embedding of the model, that expertise is a phenomenon of a society of agents not of any particular individual within it, and that all of the connotations of expertise such as cognition and learning are socially generated and socially applicable.

This paper extends the past analyses by considering how this model of expertise arises out of any modeling process applied to a persistent system that is viewed as actively participating in the processes of its own persistence, what has come to be termed an *autopoietic* system (Maturana, 1975). This analysis is intended to be neutral to issues of causality and teleology, and to present expertise and all its associated connotations as arising out of the way mere persistence is modeled. The neutrality and generality of this analysis allows neurological, behavioral, constructionist, cognitive, knowledge-level, meaning-level, and social perspectives to be reconciled, as different aspects of the phenomena of persistence and its modeling, while still providing a comprehensive account of the phenomena of expertise.

Section 2 presents a brief overview of some of the psychological perspectives applicable to the phenomena of expertise. Section 3 briefly reviews the positive feedback model of expertise formation in human society. Section 4 considers how these phenomena arise in the general case of modeling any autopoietic system. Section 5 analyzes modeling itself as a general process. Section 6 generalizes the social dimension by taking a collective stance and modeling the human species as a single autopoietic organism.

2. PSYCHOLOGICAL PERSPECTIVES ON HUMAN EXPERTISE

Much of the conceptual foundations and vocabulary of research on expert systems has arisen from the literature on human psychology, and it is worth reviewing this briefly, particularly

since it is the source of apparently quite divergent perspectives on the nature and basis of human behavior.

2.1. Cognitive Science and Behaviorism

The obvious, or most accessible, source of models of expertise for computer scientists has been the literature on cognitive psychology (Baars, 1986; Costall & Still, 1987) and cognitive science (Hirst, 1988; Meyering, 1989). This has the attraction of being based on information processing models that resemble processes familiar in the computer:

> Cognitivism is the attempt to explain human and even animal cognition in terms of internal representations and rules. The theory of perception has, for many years, been premised upon the assumption that the perceiver must construct a mental representation of the physical environment. However, technological developments such as cybernetics, information theory, signal detection theory, and, most recently, computer science, have encouraged a similar approach to the topic of 'skills' (Costall & Still, 1987).

The position of cognitive science is often stated to be antithetical to that of behaviorism:

> A revolution occurred in the 1950s which might crudely be summarized as the overthrow of Behaviourism by Information Processing (Johnson-Laird & Wason, 1977).

This viewpoint arises because the genesis of cognitive science was a reaction to statements of behaviorist methodologies that disallowed consideration of mental processes:

> Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior. Introspection forms no part of its methods, nor is the scientific value of its data dependent on the readiness with which they lend themselves to interpretation in terms of consciousness (Watson, 1913).

Skinner's "radical behaviorism" adopted a more balanced position:

> Mentalism kept attention away from the external antecedent events which might have explained behavior, by seeming to apply an alternative explanation. Methodological behaviorism did just the reverse: by dealing exclusively with external antecedent events it turned attention away from self-observation and self-knowledge. Radical behaviorism restores some kind of balance. It does not insist on truth by agreement and can therefore consider events taking place in the private world within the skin. It does not call these events unobservable, and it does not dismiss them as subjective. It simply questions the nature of the object observed and the reliability of the observations (Skinner, 1974).

From this perspective, cognitive science does not involve a revolutionary overthrow of behaviorism, but rather a redressing of the balance so that information flows, both within and without the organism, are considered as parts of an overall model. Operant conditioning is a major phenomenon in human learning and its status as a predictive model of many major phenomena cannot be overthrown. Behaviorism is itself an information-processing model focusing on the relations between sequences of inputs and outputs rather than hypothesizing

internal states, and much of cognitive science may be more accurately viewed as an extension of behaviorism.

Skinner is not concerned with cognitive models because he argues that if one is able to account for the behavior of an organism in causal terms then modeling that system as an intentional agent, while possible, is not particularly useful. His strong hypothesis is that causal modeling is always in theory possible, even though it may be impractical:

> Behavior is shaped and maintained by its consequences, but only by consequences that lie in the past. We do what we do because of what has happened, not what will happen. Unfortunately, what has happened leaves few observable traces, and why we do what we do and how likely we are to do it are therefore largely beyond the reach of introspection. Perhaps that is why, as we shall see later, behavior has so often been attributed to an initiating, originating, or creative act of will (Skinner, 1989).

2.2. Exchange Theory

Skinner's theory of operant conditioning provides an initial platform from which to develop a theory of adaptive processes in human behavior. In artificial intelligence research, reinforcement mechanisms have been incorporated as performance feedback in machine learning programs and in neural networks. Operant conditioning may also be used to model many aspects of social and organizational processes. *Exchange theory* in sociology was developed by Homans as a direct extension of behaviorism to social interactions. Some of his major propositions are:

> For all actions taken by persons, the more often a particular action is rewarded, the more likely the person is to perform that action.
>
> If in the past the occurrence of a particular stimulus, or set of stimuli, has been the occasion on which a person's action has been rewarded, then the more similar the present stimuli are to the past ones, the more likely the person is to perform the action, or some similar action, now.
>
> The more valuable to a person is the result of his action, the more likely he is to perform the action.
>
> The more often in the recent past a person has received a particular reward, the less valuable any further unit of that reward becomes for him (Homans, 1974).

These appear to be straightforward statements about operant conditioning of individuals, but Homans extends them to social behavior by bringing those who *provide* the rewards into the model, and analyzing social behavior as the *exchange* of rewards.

Homans points out in a review of the status of exchange theory in sociology that these basic behavioral propositions are universal phenomena that are adopted, often tacitly, in most other social theories (Homans, 1987). Exchange theory is not an alternative to other sociological theories, but rather a link from all aspects of social behavior to underlying psychological phenomena. It does not account for all social phenomena, but it accounts for much of their dynamics, and reminds us that an economic perspective is always available. Defining rewards, stimuli, and values are questions that may be answered in many different

ways, but if they can be identified in a particular social model then operant conditioning concepts may be applied because these concepts are inherently involved in the identification. Like Skinnerian behaviorism, exchange theory viewed in this way can be criticized for the circularity of its definitions, but such criticism misses the point that the phenomena described are ubiquitous in human behavior and in the way we perceive it—operant conditioning is both an empirical phenomenon and a natural explanatory form.

2.3. Symbolic Interactionism, Constructivism and Ethnomethodology

Viewed as a debate over the efficacy of behavioral modeling techniques, the behaviorist and cognitivist positions are complementary rather than mutually exclusive. However, both seem to offer a mechanical model of human behavior that misses a major dimension of human existence—the way in which we attach "meaning" to events and actions. Bruner saw the understanding of meaning as one of the major objectives in developing cognitive science that became lost as information processing analogs of human behavior were developed:

> Very early on, for example, emphasis began shifting from 'meaning' to 'information,' from the *construction* of meaning to the *processing* of information. These are profoundly different matters (Bruner, 1990).

The point at issue is one that has been most extensively developed by the Frankfurt School in their distinction between scientific theories and critical theories. The distinctions may be characterized through three major differences:

> Scientific theories have as their aim and goal successful manipulation of the external world. Critical theories aim at emancipation and enlightenment.
>
> Scientific theories are 'objectifying.' Critical theories are 'reflective,' or 'self-referential.'
>
> Scientific theories require empirical confirmation through observation and experiment. Critical theories are cognitively acceptable only if they survive a more complicated process of evaluation, the central part of which is a demonstration that they are 'reflectively acceptable' (Geuss, 1981).

Relative to these distinctions, both behaviorism and cognitive science are scientific theories, and neglect the way in which meaning is imposed on human behavior, becomes part of that behavior, and, like theories of behavior, comes to have major effects on behavior.

In the sociology literature, Blumer's development of *symbolic interactionism* (Blumer, 1969) addresses the role of meaning in behavior. As presented by Wallace and Wolf (1991), his three basic premises are:

> Human beings act toward things on the basis of the meanings that the things have for them.
>
> The meanings of things arise out of the social interaction one has with one's fellows.

> The meanings of things are handled in and modified through an interpretative process used by the person in dealing with the things he encounters.

What symbolic interactionism adds to exchange theory is the notion that the stimuli and rewards involved may be a matter of social convention rather than of physical phenomena, and that the conventions are not rigid but subject to an interpretive process. This is consistent with a constructivist psychology that sees individual and social constructs as generated by experience and subject to change:

> Man looks at his world through transparent templets which he creates and then attempts to fit over the realities of which the world is composed.
>
> Constructs are used for predictions of things to come, and the world keeps on rolling on and revealing these predictions to be either correct or misleading. This fact provides the basis for the revision of constructs and, eventually, of whole construct systems (Kelly, 1955).

Blumer and Kelly see meanings or constructs as being pre-linguistic, but being capable of having attached linguistic descriptions, so that language can be seen as a particular form of symbol system, and one which, as Vygotsky (1934) has stressed, can have a profound and direct influence over behavior.

Skinner would interpret meanings and constructs as hidden variables of little value in predicting and controlling behavior compared with the analysis of the preceding history of operant conditioning. Blumer and Kelly would see stimuli and rewards as meanings and constructs that themselves arise out of experience and social interaction. The behaviorist and constructivist perspectives are not in conflict. One is an external observer's view of the system with an instrumental bias, and the other is a participant's view of the system with an emancipatory bias. Where the perspectives appear to differ most is that operant conditioning is often portrayed as leading to behavioral determinism, whereas symbolic interactionism and constructivism emphasize that situations offer the choice of alternative meanings and constructions, and hence to behavioral indeterminism. In fact, as is discussed later, such indeterminism is also consistent with the dynamics of operant conditioning models.

One major source of indeterminism is the positive feedback introduced when symbolic models of behavior become part of the psychological and social systems in which the behavior arises. Garfinkel's development of ethnomethodology addresses this reflective and reflexive nature of behavior:

> Not only does commonsense knowledge portray a real society for members, but in the manner of a self-fulfilling prophecy the features of the real society are produced by persons' motivated compliance with these background expectancies (Garfinkel, 1967).

Levis has developed a theory of inter-personal behavior which provides a detailed taxonomy of the pathological phenomena that can result from our negative expectations of others behaviors in themselves coming to generate precisely those behaviors we fear. He deconstructs a wide range of Greek myths as parables providing educational examples of the social power of such interpersonal interaction (Levis, 1977).

The self-fulfilling prophecy aspect of such reflexive processes is the basis of Castoriadis' analysis of the institution of society being strictly *imaginary* (Castoriadis, 1987). We legitimate each age (Blumenberg, 1983). If this is seen to denigrate the nature of meaning then the point is being missed. It is the human capability to fantasize and reify models that are not grounded in our physical environments or biological processes that gives rise to those human characteristics that we most value. It is irrelevant to analyze these activities in terms of survival value, correspondence to reality, and so on.

3. AN EXCHANGE-THEORETIC MODEL OF EXPERTISE

An exchange-theoretic model of the social processes of adaptive systems is sufficient to account for much of the dynamics of the way in which expertise arises in society. Symbolic interactionist and constructivist analyses of the communication and modeling processes involved can then also be used to derive a cognitivist account of expertise.

3.1. Stich and Nisbett's Theory of Expertise

The basic model derives from Stich and Nisbett's theory of expertise that takes account of the sociological factors involved and is grounded in an analysis of the logic of inductive processes. They note:

> The role of experts and authorities in our cognitive lives has been all but ignored by modern epistemologists. Yet it is the hallmark of an educated and reflective person that he recognizes, consults, and defers to authority on a wide range of topics. We defer to the judgment of experts not only in assessing inference, but on factual questions as well, in medicine, science, history, and many other areas. Few educated lay persons would consider questioning the consensus of authorities on the authenticity of a painting, the cause of an airline crash, or the validity of a new theorem. Indeed, it is our suspicion that one of the principle effects of education is to socialize people to defer to cognitive authorities (Stich & Nisbett, 1984).

They analyze expertise commencing with Hume's (1739) critique of induction as a process which is not deductively valid, and where it is a circular argument to claim that it is inductively valid. Goodman proposed that we accept the circularity but note that it involves a dynamic equilibrium between data and inference rules:

> A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend (Goodman, 1973).

Rawls (1971) in his theory of formal justice as induction from informal moral behavior terms this a *reflective* equilibrium.

Stich and Nisbett argue that there are flaws in Goodman's analysis in that individuals do not necessarily behave rationally in their own behavior relative to known good practice. They repair them by proposing that the equilibrium is social not individual:

> A rule of inference is justified if it captures the reflective practice not of the person using it but of the appropriate experts in our society (Stich & Nisbett, 1984).

This gives an operational definition of the role of experts within an overall social learning process as referential subsystems managing the process in specific subdomains.

3.2. Positive Feedback Processes Generating Expertise

This analysis is remarkable in capturing the adaptive nature of expertise, socially situating it, and linking it to formal models of induction. What it does not address is how particular individuals become charged with managing the reflective equilibrium of a particular domain. Gaines (1988) has shown that a simple exchange-theoretic model can model a society of, initially identical, adaptive agents becoming functionally differentiated into 'experts' each managing learning in a particular subdomain. He notes that if tasks are themselves viewed as resources that are allocated by a client community to a problem-solving community, then there is strong positive feedback present in that clients will allocate tasks to agents on the basis of their past performance, while improved performance results from performance on tasks:

> Experts trade the skills based on their knowledge for access to problems providing them with new material for the inductive process. The formation of expertise is functional in general because it leads to division of labor in the management of knowledge acquisition. The development of an individual expert is a random process brought about by strong positive feedback loops in the social process; for example, that a proto-expert with superior performance is brought more problems and hence has a greater opportunity to learn and improve that performance. A diversity of such positive feedback processes operate in the professions and sciences with little relation between them except their overall effect in promoting the formation of expertise (Gaines, 1988).

Such positive feedback processes within the scientific community have been documented in a number of sociological studies (Blume, 1974; Hagstrom, 1965; Merton, 1973). For example, in terms of individual human expertise, a patient with a difficult condition will be routed to the specialist with the best track record in having dealt with similar conditions in the past. That specialist is thereby given the best opportunity to learn more about these conditions and hence to further enhance his or her actual and perceived expertise. Merton (1968) coined the term the "Matthew effect" for those features of the reward system in research that were biased towards allocating greater credit for the same discovery to those with an already established reputation. Similar considerations apply to the award of scholarships, invitations to scientific congresses, and so on (Blume, 1974). They also apply not only to individuals but also to social units such as a company subject to government procurement procedures that are heavily biased to contractors with 'prior experience' and with whom the government agency has 'prior experience.' Analytical and simulation models show that this positive feedback is sufficient to account for an initially uniform population becoming strongly functionally differentiated (Gaines, 1988).

3.3. The Need for Deeper Models

At this point, one might be satisfied to propose that a general theory of expertise can be developed in terms of the dynamics of a society of socially situated adaptive agents, with operant conditioning as a basic mechanism and knowledge arising as a constructed system of symbolic meanings imposed on, and becoming part of, the basic behaviors. However, it is possible to develop a broader perspective from which this theory may be derived, and where the relations between the behavioral and symbol systems, and between individuals and organizations, are more clearly defined.

Some reasons for going deeper are apparent in Stich's later repudiation of the reflective equilibrium model in his book *The Fragmentation of Reason* (Stich, 1990). He argues against it as a solution to the problem of induction rather than as a phenomenological account of expertise in society. At one level his argument can be seen as deriving from Hume's original analysis of induction as neither deductively nor inductively valid. However, Stich also argues against more specific justifications in terms, for example, of an evolutionary process that filters out agents whose behaviors are nonanticipatory for the environment in which they happen to exist. These arguments are important in suggesting that one should not be satisfied with an account of expertise that takes for granted the phenomenon of operant conditioning with the implicit assumption that the organisms involved have solved the local problem of induction by being able to adapt their behavior to anticipate their particular environment.

4. MODELING AUTOPOIETIC SYSTEMS

There is a very general systemic model of the behavior of any living system that allows the properties of socially situated adaptive agents to be derived in a way that gives new insights into, and much greater generality to, the exchange-theoretic model of expertise. In particular, one can show that neither general solutions to the problem of induction nor evolutionary functionalist arguments are necessary. Notions of agency and expertise arise out of presuppositions made in modeling living systems. They are artifacts of modeling, but highly significant ones because they are socially accepted artifacts that have become reified as part of the processes of human living systems.

4.1. Autopoietic Systems and Agency

The most fundamental properties which we impute to any system are its existence and persistence over time. A system is identifiable as not having existed before some time, of definitely existing after some later time, of persisting in existence until some later time, and of not existing again after some later time. This coming into existence, persisting for while, and going out of existence again is a common property of all systems. It applies to both living and nonliving systems, and in living systems it applies at all levels from cell to species.

What characterizes living systems are the recursive activities of self-replication underlying their persistence, that they actively and continually *create* the conditions for their persistence.

Maturana (1975) has proposed that this is the fundamental distinction between living and nonliving systems. Autopoietic systems:

> Are systems that are defined as unities as networks of production of components that (1) recursively, through their interactions, generate and realize the network that produces them; and (2) constitute in the space in which they exist, the boundaries of this network as components that participate in the realization of the network . . . a *living* system is an autopoietic system in physical space (Maturana, 1981).

It is these activities of autopoietic systems that are necessary to their persistence that we see as some form of goal-seeking and model through notions of agency.

However, there is no notion of agency in Maturana's definition, and no ascription of intrinsic intentions or goals to living systems. The notion arises because a modeler finds it convenient to adopt an *intentional stance* (Dennett, 1987), but there is no necessity to assume that the properties of the model introduced by this stance are present in the system being modeled. A reactive persistent system in itself has no goals or intentions. It reacts to its environment through mechanisms that tend to maintain its persistence despite changes in its environment. An external observer may model this behavior as goal-directed because that provides a simple predictive explanation. That is, if an autopoietic system when disturbed, regardless of what state it is triggered into, seems to return to its original state, it is naturally modeled as goal-seeking.

If the system's environment happens to contain other systems like itself, and the system's activities include observation and modeling, it may model the other systems as goal-directed, and then by analogy come to model itself as goal-directed. This is a natural outcome of autopoiesis in a social environment. The notion of agency arises as we impose notions of modeling upon the phenomena of autopoietic systems.

4.2. The Genesis of Autopoietic Systems

As well as not reading too much into models of autopoietic systems, it is important to note that we can ascribe very little to their *existence*. A chaotic universe has a probability of producing any system including autopoietic systems. Once such systems exist and are modeled, properties emerge (Sharif, 1978). Jantsch (1980) and Prigogine (1984) have developed detailed models of how organization emerges from chaos. Gould has analyzed the fossil record and modeled the genesis and extinction of a wide variety of species as low probability random events (Gould, 1989). Monod (1972) has given a biochemical model of life as an improbable phenomena that, once it exists, follows deterministic laws. When a living system comes into existence it acts to persist, but, from the systemic perspective advanced by Maturana, this is the definitional property by which we recognize its existence as a living system, not an additional property going beyond active persistence.

Barrow and Tipler (1986) have analyzed the remarkably narrow physical conditions under which life as we know it can exist—and when one examines the mechanisms by which a living organism narrows these conditions even further in order to persist, it is natural to ascribe purpose to its activity. For example, Cannon (1932) terms such activity homeostasis, and

Ashby (1952) in analyzing homeostasis as part of his *Design for a Brain*, models it as a goal-directed process. However, he also shows how such apparently goal-directed behavior arises in any system with many states of equilibrium. The utility of an intentional stance stems from simple systemic considerations, and one has to be careful in reifying the notion of agency to realize that the additional assumption of the existence of some reified 'agent' is also a matter of utility, not of existential proof or necessity.

4.3. Chaotic Behavior in Autopoietic Systems

In Ashby's day a system that reacted to its environment by acting until it arrived in a new mode of equilibrium would be seen as not only counter-acting the effects of the environment, but also arriving at some state that was determined by those effects, that is, apparently targeted upon them. Nowadays, with the realization that *strange attractors* are prevalent in all forms of physical systems (Ruelle, 1989), and particularly in biological processes and their higher-order manifestations such as brains (Basar, 1990), societies (Dendrinos & Sonis, 1990) and cultural phenomena (Hayles, 1991), it would be realized that the final state may be one of very many that have the equilibrating effect but is neither determined by the effects of the environment nor targeted upon them.

In particular, the definition of *fitness* of a species in evolutionary terms is merely a restatement of the species' persistence in terms of the environment in which it persists. As Ollason argues:

> Biologists use the concept of fitness as the explanation of the truly inexplicable The process of evolution is exactly what the etymology of the word implies: it is an unfolding, an indeterminate, and in principle, inexplicable unfolding (Ollason, 1991).

A species is fit to exist in an environment in which it happens to persist. As noted in the previous paragraph, this does not mean it was targeted on that environment or that there is a determinate relation between the nature of the environment and the species that happens to have evolved. The environment acts as a filter of species and those that persist are fit to survive. There are no teleological implications, nor moral implications that the species has more right to survive. Neither does this model give survival-directed behavior any greater probability of leading to persistence than any other behavior. Gould (1989) details the random phenomena that have made particular species fit to persist for a while in the fossil record. Bickerton (1990) argues that there is no evidence for what we deem to be high-level human traits to have survival value—intelligence and language have at least as many disadvantages as advantages, and may be seen as of negative value to the survival of the human species.

Thus, in adopting an intentional stance it is important to remember that one is selecting a modeling schema for its simplicity, convenience, and utility. Notions of expertise, the connotations that come with them, and the psychological and social behaviors we associate with expertise are part of a modeling schema that we impose upon certain phenomena of our species. It is not only relevant to ask what is this modeling schema and how does it operate, but also why is it simple, convenient, and useful. One can then also extrapolate this line of thought and propose other modeling schema that may be even simpler and more convenient and useful.

4.4. Models of Autopoietic Systems

In explicating modeling one becomes involved in a reflexive process because one has to state how the notions of modeling arise from the models adopted of autopoietic systems. However, the notions involved are naturally recursive, and involve the same logic of existential assumptions as in the previous section—one has to assume the existence of modeling processes before the notion of modeling itself can be analyzed.

In modeling the persistence of an autopoietic system despite changes in its environment, one finds it convenient to classify the unobservable state of the system by the observable state of the environment. That is, even though the actual state of the system is not determined by the environment, it is convenient for purposes of explanation to group together internal states by their effects in enabling the system to cope with particular states of the environment. Given that the system persists, as its environment changes, it will be seen as generating a trajectory of states each one of which is indexed by the state of the environment. It is in this sense that an observer imputes to the system a 'model' of its environment. This is the argument form of the result in cybernetics that "every good regulator of a system must be a model of that system" (Conant & Ashby, 1970). The classification of the system's state in terms of that of its environment also makes an 'intentional stance' self-justifying. The system, if it persists, by definition comes to be in a state that compensates for that of its environment.

As a concrete example, consider the process whereby sand takes up the shape of an irregular container. One can examine the end result and say that the sand has 'modeled' the container. However, one would find no modeling algorithm in the grains of sand, and the existence of the model itself is a construct of the observation that the sand and the container now match one another globally—it is only in a very remote sense a property of the sand. If one wished to expedite and improve the process—to *educate* the sand—one would be foolish to attempt to aid the local modeling process by moving individual grains of sand. Instead, one would shake the container to provide an environment of maximum opportunity for the sand to settle into the shape of the container. One would shake roughly at first and then increasingly gently. This is exactly the *modeling* process of the Boltzmann machine and its management through simulated annealing (Aarts, 1989). It also corresponds to Gibson's (1979) *ecological* model of perception in which we model the world by the brain coming into equilibrium with the stream of incoming sensation.

The theoretical foundations for the relation between an autopoietic system and its apparent modeling capability can also be given a simple interpretation. If we consider a perfect system that is able to match every change in the environment, then an observer modeling each of the two systems through a behavior-structure adjunction will ascribe the same minimal structure to both up to an isomorphism (Gaines, 1977). Because the observer obtains the same model of both system and environment, it is a reasonable metaphor that the autopoietic system contains such a model of its environment. For example, an operator controlling a linear mechanism apparently has an internal linear model, despite the discrete 'bang-bang' nature of the neuromuscular process, precisely to the extent that his or her control is optimal (Gaines, 1969).

The symmetry of the models derived by an external observer of a coupled system and environment may appear paradoxical. If the models obtained are the same, then it makes as much sense to see the environment containing a model of the system as it does to see the system having one of the environment. The symmetry can only be broken by presuppositions about the nature of the coupled systems, or by additional information about one or both of them. If the observer has access to other sources of information about one of the systems that enables its behavior to be modeled as causal dynamics responding to observable inputs, that model will be preferred to the less predictive one of autonomous agency. If such information is available to others but a particular observer cannot access it and model the environment causally, and hence takes an intentional stance, then more informed observers may find this inappropriate and term it unnecessary *animism*. A behaviorist will see an intentional stance as animism even when applied to a person.

4.5. Luhmann's Systemic Sociology

Luhmann's systemic sociology provides a very detailed analysis of the way in which socially meaningful behavior arises out of the dynamics of autopoietic systems. He commences his analysis of the symbolic concepts of trust and power with a simple behavioral model and its interpretation:

> The world is overwhelmingly complex for every kind of real system . . . Its possibilities exceed those to which the system has the capacity to respond. A system locates itself in a selectively constituted 'environment' and will disintegrate in the case of disjunction between environment and 'world.' Human beings, however, and they alone, are conscious of the world's complexity and therefore of the possibility of selecting their environment—something which poses fundamental questions of self-preservation. Man has the capacity to comprehend the world, can see alternatives, possibilities, can realize his own ignorance, and can perceive himself as one who must make decisions.
>
> . . . we invoke a whole new dimension of complexity: the subjective 'I-ness' of other human beings which we experience (perceive) and understand. Since other people have their own first-hand access to the world and can experience things differently they may consequently be a source of profound insecurity for me (Luhmann, 1979, p. 6).

Crespi criticizes Luhmann's statement in terms of it being based on an intentional stance that is not made explicit:

> The concept of complexity reduction which forms the basis of Luhmann's theory implies reference to an active principle of selection, which can be understood only as the negative capacity of consciousness in its relation to objectivated forms. Luhmann, even though he recognizes the importance of negation in the concept of complexity reduction, does not relate the latter to subjectivity but to the social system itself, thus attributing to the system a quality that can only be specific of an individual (Crespi, 1989, p. 14).

However, chaos is maximally complex and complexity reduction corresponds to the emergence of organization. The logic of Luhmann's statement can be reformulated to avoid any apparent imputation of intentionality. A chance fluctuation in complexity is sufficient

to account for the emergence of the system-environment relations he specifies. Thereafter, modeling and multi-level reflexive modeling lead to the phenomena he describes, including those that Crespi ascribes to the subjectivity of conscious individuals. The essence of the arguments in this section is that there is no conflict between Crespi's and Luhmann's positions, both of which are developed in their works into rich analyses of significant phenomena in the social behavior of living systems. This compatibility is apparent in Crespi's affirmative quotation of Touraine's analysis of the autonomy of action:

> The sense of an action can neither be reduced to the adaptation of the actor to a more or less institutionalized normative system, nor to the expression of the mind through a social activity (Touraine, 1965).

The actions of an autopoietic system are prior to, and independent of, such considerations, which belong only to the interpretive act of modeling the system. They are an interpretation, not a derivation, of its behavior.

5. CONSTRUCTS OF MODELING

The preceding section has argued that the behavioral phenomena of an exchange-theoretic model of expertise derive from the presuppositions made in modeling autopoietic systems, in particular, the mutual models of interacting autopoietic systems. This section addresses the other aspect of cognitive modeling of such systems, how notions of meaning arise out of the modeling process. It does so by deconstructing an abstract model of general modeling processes. The argument is again a reflexive one, that it is the way in which we model modeling that leads to our notions of higher-level cognitive processes.

5.1. Modeling Modeling

Klir (1976, 1985) has analyzed the processes involved in any modeling system and developed an infrastructure for them that can be instantiated in different ways to encompass many different modeling schema. His basic constructs form a hierarchy of systems: a *source system* providing descriptive terms, a *data system* providing descriptions in these terms, a *generative system* providing a regeneration of these descriptions in terms of a *structure system* providing theoretical terms, itself described through *meta-systems, meta-meta-systems,* etc. Figure 1 shows this modeling hierarchy expressed in terms of a primitive process of forming a construct by making a distinction (Gaines & Shaw, 1984). Thus, a general modeling system may itself be modeled as a process that makes distinctions in the world, gathers data in terms of those distinctions, selects from a repertoire of representations those which best generate the data, analyzes relations between the structures of such representations, and recursively repeats such analysis to generate higher levels of the hierarchy. The term *anticipation* is used for the output to capture both prediction and action. It is not necessary in general to distinguish whether the system anticipates correctly by passive prediction, or by actively changing the world to be predictable.

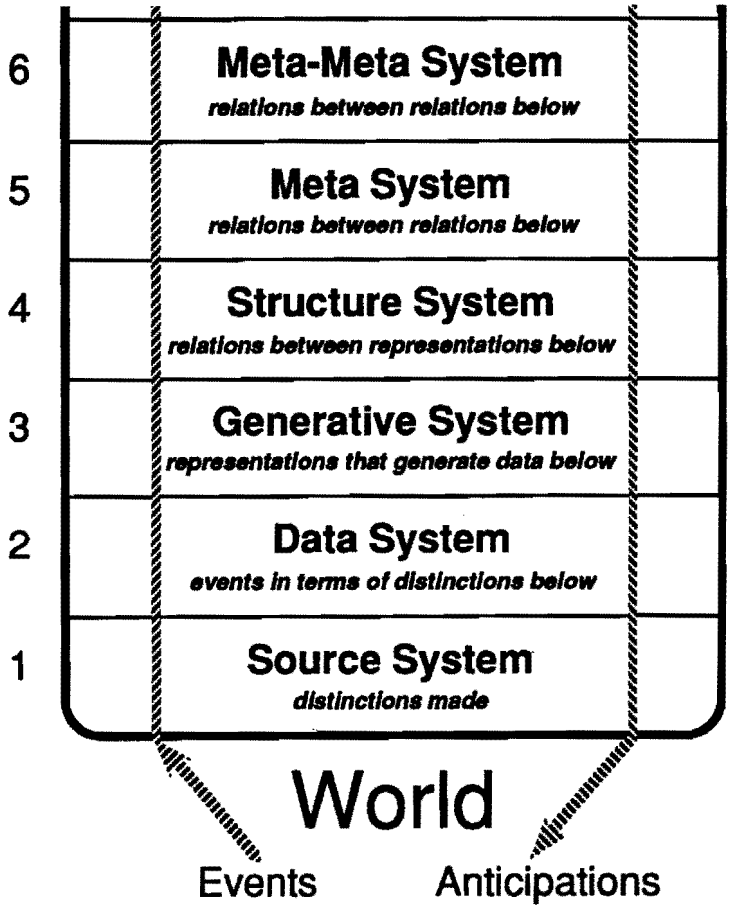

Figure 1. Klir's modeling hierarchy.

The instantiation of this hierarchy with particular data description languages, model classes, and measures of model complexity and model-data approximation has proven in practice to account for the wide range of mathematical modeling techniques (Gaines, 1977; Klir, 1985). It is interesting as an account of generalized adaptive processes since it makes clear the presuppositions necessary for modeling to take place—a *tabula rasa* cannot begin to model and some degree of 'innateness' is required. The trade-off between the amount of data needed in learning and the innately 'assumed' constraints upon the world can be investigated (Gaines, 1976), as can that with the sociocultural filtering of data to improve its support of learning (Gaines, 1989). Fodor-style arguments about the innateness of cognitive processes (Fodor, 1986) can be investigated in terms of the presupposed structures at each level of the modeling hierarchy, together with the possible cultural mechanisms for

ensuring that the presuppositions are correct and met in experience. In essence, one can operationalize Kant's *synthetic a priori*, the phenomenon where we bring to our apprehension of the world presuppositions that do not become effective until they are used to model experience:

> Though all our knowledge begins with experience, it by no means follows, that all arises out of experience. For, on the contrary, it is quite possible that our empirical knowledge is a compound of that which we receive through impressions, and that which the faculty of cognition supplies from itself (Kant, 1781).

Pope (1984) has noted that distinction-making hierarchies such as that of Figure 1 are common in a wide range of different cultures, and that they have another dimension in that the recursive upward-building process may be seen for what it is and transcended. We may interact with a world, begin to make distinctions about it, begin to make distinctions about these distinctions, and then recognize the distinction-making process. This does not necessarily prevent us from continuing to build the hierarchy but it does also allow us to step outside it and reflectively treat our own modeling processes as phenomena which we are able to model. This reflective step, of 'modeling the modeling' is distinct from the abstractive set of subsuming more and more modeling within a unified framework.

If the world being modeled is a sociocultural world of other modeling agents there is an intrinsic instability in the model process. My model of you is affected by your model of me, and by my model of the modeling processes underlying your model of me, and by my model of your model of the modeling processes underlying my model of you, and so on. Lefebvre uses this mutually referential framework to develop a formal basis for ethical systems based on remarkably sparse logical primitives (Lefebvre, 1982). It is also the basis for ethnomethodological analyses of the essential reflexivity of society.

5.2. Cognitive Processes in Modeling

One may instantiate the modeling hierarchy in a number of different contexts using the vocabularies of psychology, anthropology, organizational science, education, artificial intelligence, and so on, to develop systemic architectures for modeling processes in a wide range of systems. Figure 2 shows a variant of this hierarchy emphasizing the cognitive processes involved in modeling the world and their definitions in abstract, systemic terms (Gaines, 1989):

- To *recognize* at the lowest level is the capability to notice recurrence of the 'same' events in the world when they recur. This is already a significant cognitive act because 'same' is subject to personal definition, and the concept that events recur is a strong presupposition. Recognition is fundamental to any modeling system but is in itself a weak operation since it is dependent on the recurrence to make use of the data.
- To *recall* at the next level is the capability to regenerate the distinction used in recognition internally so that it is itself an 'event' that may be processed. This facility

to recreate events in the 'imagination' is fundamental to the existence of the cognitive process, detaching human knowledge processing from the immediacy of experience.

- To *represent* at the next level is the capability to derive the distinction used in recognition and recall from other distinctions that may themselves not relate directly to experience. This facility to 'represent' events in terms of distinctions that relate only indirectly to experience is again fundamental to the efficiency of the cognitive process, allowing novel distinctions to be developed that efficiently encode wide ranges of otherwise unrelated experience.
- To *reconstrue* at the next level is the capability to derive one distinction from multiple models. This facility to move between modeling systems is fundamental to the adaptability of the cognitive process, and the human species, allowing a wide repertoire of anticipatory subsystems to be developed to cope with the variety of the world.
- To *abstract* at the next level is the capability to detach distinctions from their sources and make the relations between models themselves subject to study. This facility to study the world of modeling as if it were a world of experience is fundamental to the externalization and growth of human knowledge as a cumulative by-product of the anticipatory process. It makes the expertise of the species largely independent of that of existing individuals.
- To *originate* at the next level is the capability to treat the distinction making process itself as a human activity, subject to choice and change, and to generate distinctions in themselves. This freedom raises many questions as to the nature of 'reality,' of the 'wisdom' of certain distinctions, and of the relationship of the distinction-making, cognitive, and knowledge processes to the nature of the individual and the survival of the species.

5.3. Information Flows in Modeling

The information flows in the hierarchies of Figures 1 and 2 are very significant. In anticipation the key distinction is to what degree a level accounts for the information flowing through it and hence this distinction may be termed one of *surprise* (Gaines, 1977). Surprise goes in opposition to the degree of membership of an anticipated event to an actual event and the expected surprise is a form of entropy. Surprise at the lowest level of the hierarchy corresponds to distinctions being inadequate to capture events; surprise at the next level to inadequate variety to experience events; at the next level to inadequate approximation to predict events; at the next level to inadequate simplicity to explain events; and at the next level to inadequate comprehensiveness to account for events.

The formal theory of modeling is one in which models are selected at each level down the hierarchy to minimize the rate at which surprise is passing up the hierarchy. The criteria for model selection independent of the data are generally thought of as being ones of simplicity/complexity: of two models which fit the data equally well, choose the simplest. However, notions of simplicity/complexity are not well-defined nor intrinsic to the class of models. The simplicity/complexity ordering is arbitrary and in its most general form is just one of preference. In situations that are mathematically well-defined, such as determining the

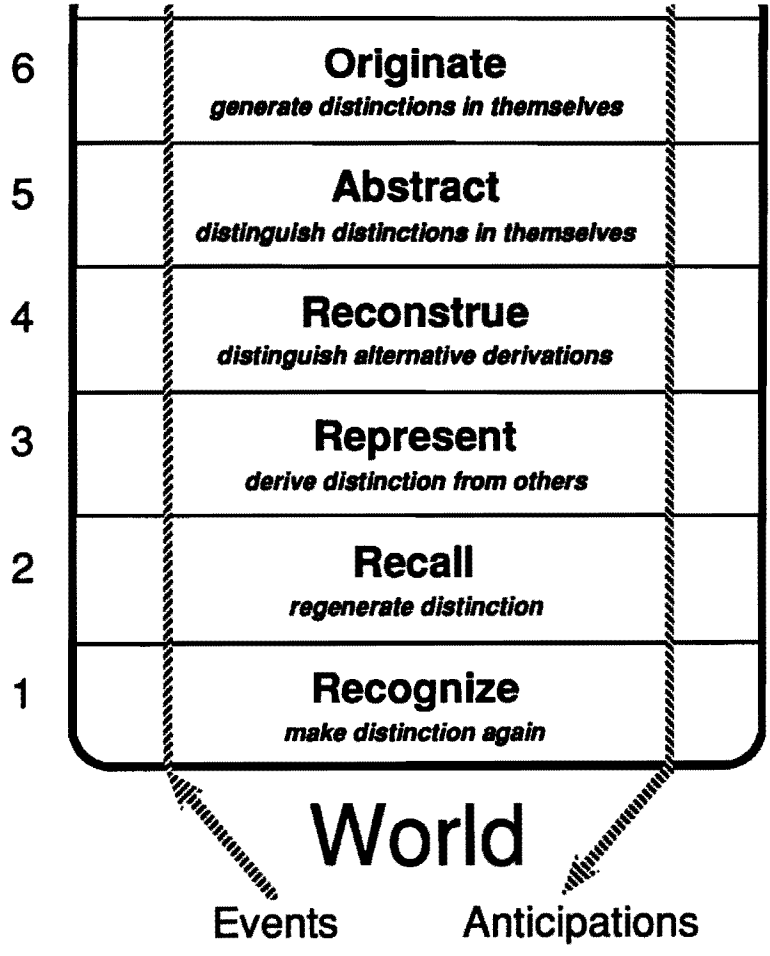

Figure 2. The modeling hierarchy in a psychological instantiation.

structure of a stochastic automaton from its behavior, such a model schema gives the correct results (Gaines, 1977). Conversely, the success of the schema in stabilizing with regard to a given world defines the characteristics of that world.

Thus the basic modeling schema for adaptive behavior is one in which surprise flows up the hierarchy and preferences flow down. In primitive organisms only the lower levels of the hierarchy are developed, surprise and preferences are linked directly to experience, their mechanisms are genetically determined. In higher organisms the modeling process generalizes both surprise and preference to cope with novel environments. Human life has developed

the upper levels of the hierarchy and detached surprise and preference from direct experience and from their genetic roots. Surprise can flow up from a level without flowing into it from below because the processes at that level have generated novelty. Preference can be generated at a high level detached from experience and flow down to affect the relations of the organism to the world.

There is physiological and behavioral evidence of the existence within the brain of the two channels of communication shown in Figure 2 (Tucker & Williamson, 1984). The *arousal* system passes surprise upwards to the cortex from the limbic region when unexpected events occur. The *activation* system passes preferences down from the cortex to the motor regions. Emotion also has a direct interpretation in terms of these information flows. Melges notes that:

> The normal function of emotions is to attune the person to overall discrepancies between the present and the future so that he adjusts his plans of action to his future images (Melges, 1982).

Thus emotions may be seen as derived from surprise with the type of emotion varying according to circumstances. The deviation from the model may be construed as having adverse or beneficial consequences, being distracting, requiring attention, investigation, action, and so on. This is consistent with Kelly's notion that negative emotions arise through the violation of core constructs (McCoy, 1981). In the modeling hierarchy such core constructs are distinctions that we prefer not to change. From a systemic point of view human feeling tones are signals directing the inductive inference process. This is the basis of Gray's (1979) insight that emotions structure cognition and memory—a modeling system does not need to store what it can anticipate, and hence changes in a system's model that we interpret as 'memory' processes correspond to failures in anticipation which are associated with emotions that are the system's cognitive interpretation of surprise.

The analysis of modeling and its infrastructure have generated a perspective in which the emphasis on information flows in cognitive science is natural and expected. However, there is nothing that has been assumed that differentiates the modeling processes of the person from those of any other cognitive system. Notions of autopoiesis and modeling apply as well to organizations as to people. Even notions of emotion are applicable to organizations, not as metaphors, but as an interpretation of the evaluation of discrepancies in the modeling process. The following section generalizes the analysis of autopoietic systems and modeling developed so far to apply also to organizations.

6. HUMANITY AS AN AUTOPOIETIC SYSTEM

The analysis to this point has been couched largely in terms of individual human agents because this gives the concepts involved in a conventional grounding and allows them to be evaluated in terms of familiar intuitions. However, the interactions between individuals involved in the exchange-theoretic model of expertise has been treated as a different construct from the information flow within an individual, and organizations have played no essential part in the discussion. The final step in the argument is to change perspectives once again and adopt a *collective stance* which considers the human species as a single autopoietic

system. This allows a unified treatment of individuals and organizations, and of information flows within individuals, between individuals, and between individuals and organizations. The phrase collective stance is chosen by analogy with Dennett's (1987) intentional stance, because its primary justification is one of utility.

6.1. The Human Species as a Single Organism

To allow the social dimension of expertise to be treated as a integral feature of human activity, it is convenient to consider humanity as a single organism. This is intended as a perspective rather than a deep ontological commitment. It does not involve teleological issues such as the ascription of agency or goals to the species, because neither agency nor goals will be treated as causal variables for any organism. They, and many other related descriptive terms, will be treated as terms within modeling vocabularies, but no particular class of models will be taken as primal. The same considerations apply to notions such as causality. The universe is taken to be something which exists and exhibits phenomena, and the initial descriptions from which notions of expertise are derived are intended to be read phenomenologically. That is not to say that they are intended to be theory-free, but rather that they are about *what happens,* not how or why.

How humanity came to be regardable as a single organism is not a major issue for this paper. One could assume that it has always been so, and that we tend not to see it so because egocentric perspectives emphasizing individual humans are more useful in our everyday lives and cultures.

One could assume that it was not always so, but that humanity has evolved to be more naturally and usefully regarded as a single organism. Again, such considerations involve no deep issues since no such assumptions are necessary to the argument, and both these perspectives and their associated vocabularies will be treated as legitimate models. In particular, notions such as fitness will not be taken as explanatory variables but rather as post-hoc rationalizations.

The basic mechanism of evolution is that organisms change in ways that are influenced by its environment. The phenomenon is one merely of change, not of direction. If it is modeled as a process of improvement then this modeling process may itself be modeled as one of defining what is meant by improvement. From a phenomenological point of view, the way we explain the universe involves circular arguments, and where we chose to break the circularity is arbitrary in itself but significant in leading to our world-views. Ultimately the world-views that we happen to have are just that—the world-views that we happen to have. If we wish to derive them from arguments relating to their appropriateness to the universe that we happen to exist in, then those arguments presuppose a world-view that we happen to have that lets us do so. The circularity is inescapable but it does not make the whole process of modeling meaningless. Rather, it gives us a choice of perspectives and the basis for changing perspectives should we chose, and the possibility of persuading others that the change is worthwhile in terms of some world-view.

6.2. Recursive Partitioning of Humanity in Time and Space

The organism that is humanity will be assumed to exhibit certain phenomena.

1. That it may be recursively partitioned to some depth, and that the parts exhibit much the same phenomena as the whole.
2. That the parts and the whole are distributed in space and have limited extents in space and time.
3. That the parts and the whole interact with environments which may include the parts or the whole.
4. That the parts, the whole, and the environment change as they interact.
5. That some such changes may be evaluated as improvements.

What is being defined here is a vocabulary for describing phenomena of an organism that may be regarded as partitioned into a variety of suborganisms interacting with environments and exhibiting performance change. If this vocabulary already appears to involve strong presuppositions and to be theory-laden, then three comments may be applicable.

1. It is very difficult to develop any description of humanity encompassing normal models without the presupposition that such a descriptive vocabulary may be used.
2. If these presuppositions are taken as themselves being derivable then an alternative framework is being proposed. For the purposes of this paper, the terms are being presupposed as applicable primitive descriptions.
3. The questions of what are space and time, how the partitioning is done, what are the environment, interaction, and change, and how improvement is evaluated, are all as significant as they may appear. However, it is not necessary to answer any of these questions to use the framework. That is, the answers are arbitrary for the purposes of this paper. Differing answers to questions such as these, however, can generate apparently very different world views.

Pepper (1942) analyzed such different world views as *world hypotheses*, defined as unrestricted products of knowledge which cannot reject anything as irrelevant. He proposes that someone desiring to understand the world will take an area of commonsense fact and use this as a basic analogy with which to understand other areas. The structural characteristics of this *root metaphor* become the categories for that person's basic concepts of explanation and description. Pepper considers four root metaphors underlying common world hypotheses: *formism, mechanism, contextualism,* and *organicism.* However, these are only some of an infinite range of possibilities. From a systemic point of view, world hypotheses are complete, autonomous, noncomparable, and can completely account for one another. They are cognitive perspectives on the world, not competing theories. If they compete it is in their attractiveness and utility, not in their truth.

6.3. Expertise as a Positive Feedback Phenomenon

The overall thesis developed in this paper is that the systemic mechanism underlying the phenomena of expertise is a simple positive feedback process within an autopoietic system that generates complex behavior which is subject to many interpretations according to the parts and levels of the systems involved. An account of expertise involves first giving a concise description of the systemic mechanism and then examining how, and why, it may be interpreted in many different ways. That is, what are the presuppositions underlying different perspectives, what models of expertise arise given these presuppositions, and why is each set of presuppositions reasonable in some sense? While this approach is integrative, there is no reason to suppose that the differing perspectives will lead to commensurable accounts, or a uniform theory applicable to all phenomena. In fact, the contrary will be argued, that the essential characteristics of expertise are those of a variety-generating mechanism. From a functionalist perspective, such a mechanism supplies an organism with a reservoir of behaviors with which to cope with an uncertain environment.

Notions of expertise arise because the organism adapts as a whole through adaptation of its interacting parts. The phenomena of expertise precisely correspond to those leading to distribution of tasks and functional differentiation of the parts. The mechanism is one of positive feedback from parts of the organism allocating resources for action to other parts on the basis of those latter parts' past performance of similar activities. Distribution and differentiation follow if performance is rewarded, and would-be performers of tasks, being excluded by the feedback mechanism from opportunities for performance of those tasks, seek out alternative tasks where there is less competition.

The organism thus conceived has a potential for adaptivity through functional evolution and differentiation that derives from changing relationships between its parts rather than changes in their biophysical substrate. No one part is crucial to the functioning of the organism, but the experience of any part can contribute to the adaptivity of the whole because of the discourse between parts. In other words, the system can be said to support *cultural evolution* (Boyd & Richerson, 1985) through a variety of *memetic* (Dawkins, 1982), rather than genetic, processes. A major difference between cultural and genetic evolution is that parts which fail in genetic evolution only contribute indirectly to the success of others, whereas those that fail, even to destruction, in cultural evolution can communicate information that contributes directly to the success of others.

The phenomena of expertise correspond to a society of relatively undifferentiated adaptive agents becoming functionally differentiated through decentralized systemic processes. From a functionalist perspective, many other phenomena may be analyzed as byproducts of this process. The partitioning in time corresponds to the limited life and fragility of the parts, and the coordination of the parts over space and time gives rise to phenomena of discourse, language, and overtly represented knowledge. An organism structured as a society of agents seems well-fitted to an ecological niche involving a greater variety of conditions and greater changes in conditions than those for other known animal species. In particular, it may be capable for some time of supporting a growth in biomass that is very high in relation to the available natural resources, and such that the growth itself is the primary cause of radical changes in the environment of the species.

The limited capacity for activity and limited life of the agents in whom expertise is focused makes it desirable to encourage them to replicate that expertise by any means possible. Hence the dissemination of expertise is itself rewarded and social mechanisms such as apprenticeship, and oral and written communication of the basis of expertise are developed. The general increase in levels of specialist expertise as the organism as a whole adapts, means that new parts have an increasingly long trajectory of adaptation to become competitive, leading to social mechanisms such as universal education.

6.4. Neural Structures, Individuals, Groups, and Organizations as Different Partitions

The model developed can be applied integratively to analyze the relations between levels. If we return to the species perspective for a moment and consider the cognitive unit to be the biomass, then in terms of territory, diversity, and risk it makes sense for that biomass to distribute itself in such a way that severe damage to its parts has little effect on its whole, yet models generated by its parts may be shared by the whole—a formula for distributing the whole brain among small mobile sculls. Our folk evolutionary biology may tell us that is not how it happened, but our systemic analysis will tell us that it does not matter too much if our model is not historically sound provided it is a good model of the system now. An analysis of the psychological processes of individuation arising out of the extreme dependence of the child on the mother and the resultant primate nurture practices may suggest that our naive model is not so wrong after all.

The simple model that the distributed intelligence making up the biomass has people as functional units, however, misses much of the richness of the processes involved. We are egocentric and have somehow located our egos in our skulls. However, from a systemic point of view any cross section of the modeling hierarchy may be a possible partitioning of the agents in a distributed cognizant system, whether organization, group, person, role, module, or neural complex. Simmel made this interrelation of wholes and parts the central theme of his sociology:

> Society strives to be a whole, an organic unit of which the individuals must be mere members. Society asks of the individual that he employ all his strength in the service of the special functions which he has to exercise as a member of it; that he so modify himself as to become the most suitable vehicle for this function. . . . man has the capacity to decompose himself into parts and to feel any one of these as his proper self (Simmel, 1950, pp. 58–59).

Simmel's insight that the group member is always a fragment of a person, a role created precisely to enable the person to enter the group has been developed extensively by Wolff with his notions of *surrender* and *catch*:

> From the standpoint of the world of everyday life, the mathematician, as we often put it, lives in the 'world of mathematics', dealing with 'nonreal' elements, notably numbers, whose relation to 'real' things, to 'reality', is not part of his concern. Analogously for the logician. What makes our subject-object approach to this attitude misleading is the fact that the subject, the student of mathematics or logic—his or her individuality, including motives and attitudes—is irrelevant for

> our understanding; the only thing that counts is the pursuit, with its results and questions (Wolff, 1976, pp. 162–163).

He makes the key point that not only does the real world, the object, disappear to be replaced by the world of mathematics, but also that in entering into this world the person doing mathematics, the subject, also disappears to be replaced by a new entity, the mathematician.

Wolff was not alone in these insights and there are two further models which develop and complement his. In terms of his *subject,* Pask's concept of *P-individuals* as coherent psychological processes capable of engaging in conversations (Pask, 1975) is a useful representation of the results of Wolff's *surrender*, particularly when we note that several *P-individuals* may execute within a single processor. Thus 'the mathematician' may engage in conversation with 'the physicist' or 'the statesman' all of whom happen to use the same brain for their processing. In terms of Wolff's *object,* Popper's concept of a *third world* of *statements in themselves* (Popper, 1968) is a useful representation of that which we *catch,* emphasizing the distinct ontological status of Dawkin's memes:

> I regard the third world as being essentially the product of the human mind. It is we who create third-world objects. That these objects have their own inherent or autonomous laws which create unintended and unforeseeable consequences is only one instance (although a very interesting one) of a more general rule, the rule that all our actions have such consequences (Popper, 1974, p. 148).

Pask goes beyond Simmel in conceiving that a *P-Individual* is not just what we conventionally term a role within a person (which he terms an *M-Individual*, or mechanically characterized individual), but may itself be composed of a number of roles coming together to form a unity that we conventionally term a group or organization:

> A P-Individual . . . has many of the properties ascribed by anthropologists to a role, in society or industry, for example. A P-Individual is also a procedure and, as such, is run or executed in some M-Individual, qua processor. However, it is quite exceptional to discover the (usually assumed) one to one correspondence between M-Individuals and P-Individuals (Pask, 1975, p. 302).

Shaw has developed Pask's notions to show how personal construct psychology may be used to account for the psychological processes not only of individual people but also for that of functional groups such as a nuclear family or a product executive (Shaw, 1985).

Thus, the human species as variously distributed intelligence may be analyzed within the framework of the modeling hierarchy already developed. Figure 3 shows how the species itself may be treated as an anticipatory system with various imputed modeling processes that may be seen as carried out by various levels of suborganization, themselves constituted through roles within individuals. From this perspective organizations are cross-sections of the representation of the overall anticipatory system within the modeling hierarchy, and roles are cross-sections of organizations within the same framework. Communication and coordination processes occur at every level of the hierarchy, and can be analyzed in terms of the forms of representation and activity defined at each level.

In particular, this analysis suggests that theories of cognitive agencies should be applicable to all levels of intelligent organization, not just the human person. It draws attention to the

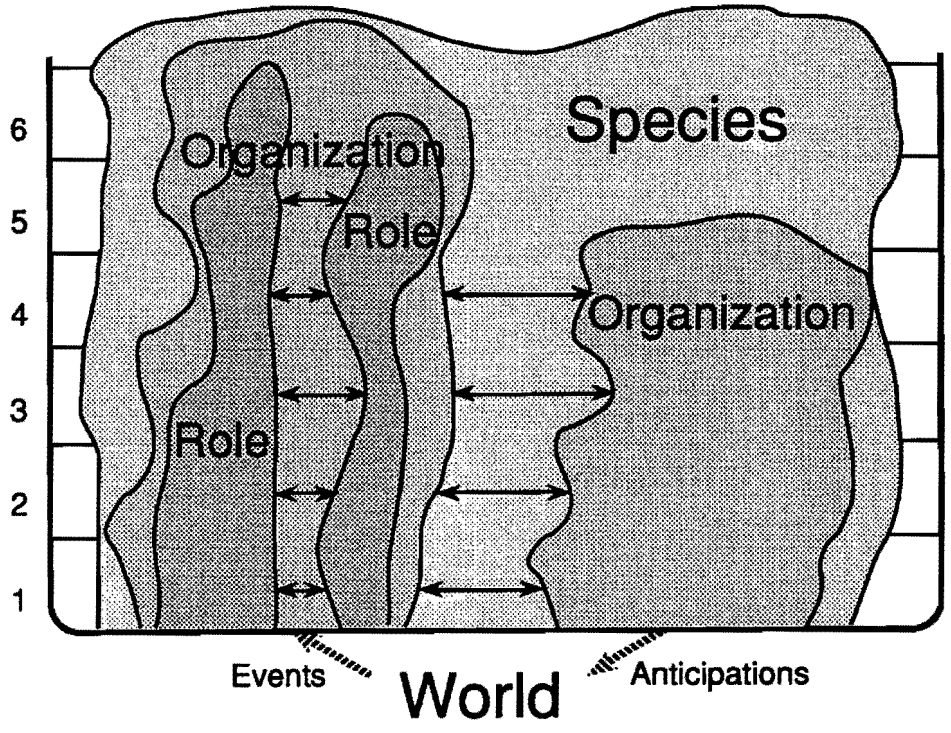

Figure 3. Psychological and social organizations as cross sections of the modeling hierarchy.

potential artifacts inherent in perspectives that place undue emphasis upon the individual, reifying one aspect of localized cognitive activity and neglecting its embedding in a larger context. Figure 3 emphasizes that the boundaries placed around cognitive entities are themselves part of the process whereby we model them, and different placements will lead to different, but related, models. When we study the intermodel relations in cultural, organizational, and social psychology, it is illuminating to see them generated by distinctions that are themselves part of the processes being studied. We cannot escape the modeling process, but we can attempt to make overt its current preconceptions, deconstructing not just our theoretical frameworks, but the personal, cultural, and social environments in which they have been developed.

6.5. Language and Knowledge

We can gain new insights into the nature of language and knowledge within this framework of Figure 3. If the biomass is partitioned into semiautonomous, physically distinct units, how

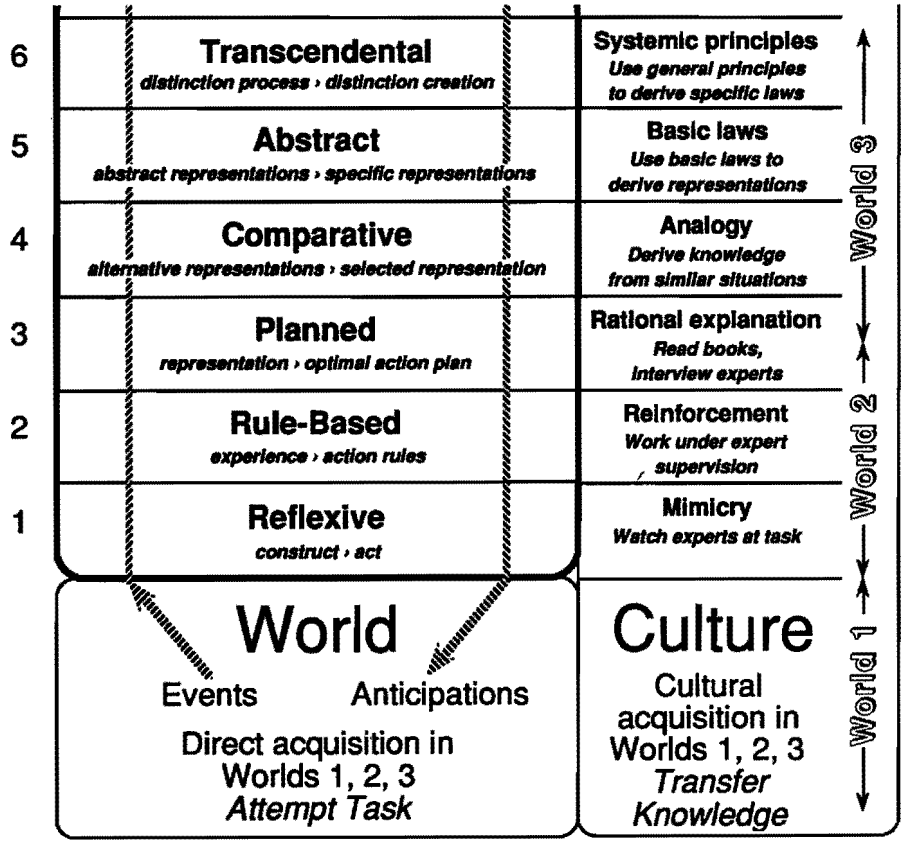

Figure 4. Control and coordination in the modeling hierarchy.

does it coordinate the activities of its components? Systemically, language may be seen as a byproduct of such communication. What functions language has in a particular situation will depend on what levels within the hierarchy coordination is required. In ethnomethodological terms, language and knowledge have to be seen as *situated* in the total context of activities within the lifeworld (Garfinkel, 1967).

At the lowest levels of the modeling hierarchy it is not the symbolic aspect of language which is relevant. We coordinate reflexive actions at the lowest level by mimicry. We coordinate rule-based conformity at the next level by reinforcement. These are the levels of Hall's (1959) *informal* and *formal cultures* respectively. It is only at the next level that his *technical culture* and overt knowledge constructs comes into play. As we go up the hierarchy our linguistic needs become more abstract and what is communicated becomes less associated with experience and more with the way in which we communicate our models in themselves. These are the levels of Popper's third world of objective knowledge. Figure 4

attempts to capture these concepts by showing the modeling hierarchy on the left in terms of connections between events and anticipations at each level, and on the right the coordination, communication, and knowledge transfer mechanisms involved across society.

Thus language and knowledge can be seen as byproducts of distributed agency. The work of Vygotsky (1934) and Luria (1961) on the role of language in aiding the performance of perceptual-motor skills fits within this framework as communication between two roles, or Paskian *P-Individuals,* within one head. However, the framework bypasses, as I think it should, the role of language in the mental processes of the individual. Language may be needed for internal communication between Simmel's fragments that we have created in order to play roles in society. Because it is needed for external coordination it may well become used for internal coordination. However, there is no systemic reason for this to happen universally. Most of our mental processes may be not so much prelinguistic as utterly independent of language. This is a matter for empirical study.

7. CONCLUSIONS

The main outcome of this paper is a recommendation to adopt a *collective stance* to humanity and see it as a single organism, a neural network that is distributed in time and space by recursive partitioning into parts similar to the whole.

The phrase *collective stance* is chosen by analogy with Dennett's *intentional stance*, because its primary justification is one of utility. A collective stance provides a convenient perspective from which to view phenomena of human existence, including behavioral and knowledge processes, and the specification, design, application, and impact of technological support systems.

The parts into which the human organism is recursively partitioned include societies, organizations, groups, individuals, roles, and neurological functions. Many concepts that apply to individuals may be applied to social systems without recourse to metaphor or analogy for systemically they are the same.

Notions of expertise arise because the organism adapts as a whole through adaptation of its interacting parts. The behavioral mechanism is one of exchange of reinforcement through some parts allocating tasks to others. The preferential allocation of tasks to those parts which have performed well previously also gives those parts access to experience enabling them to adapt and perform better in the future. This positive feedback leads to functional differentiation of the parts and the distribution of activities.

From a functionalist perspective, the coordination of the activities among the parts leads to phenomena of communication, discourse, and language. The short lives of individual parts would lead to loss of knowledge by the organism as a whole unless compensated by social interactions supporting knowledge transfer, including the generation and storage of overtly represented knowledge.

The improved performance resulting from adaptation may be modeled as the part involved having acquired a model of its task. Reflective processes in which parts model the behavior of other parts, including themselves, leads to a hierarchy of models of increasing abstraction, and detachment from direct experience.

The modeling of human activity in terms of behavioral contingencies and its modeling in terms of symbolic interactions are complementary analyses of the same phenomena. There will be some situations which are more richly represented by one of the analyses and poorly by the other. Many of the phenomena of human action and expertise are behavioral and do not involve significant symbolic representations.

The Copernican revolution greatly simplified our model of the solar system by dropping an earth-centered stance and adopting a sun-centered stance—we ceased to place ourselves at the center of the physical universe. The call for a collective stance proposes a similar revolution that simplifies our model of the human system by dropping an individual-centered stance and adopting a collective stance—we cease to place our individual selves at the center of the cognitive universe.

ACKNOWLEDGMENTS

This work was funded in part by the Natural Sciences and Engineering Research Council of Canada. The arguments in this paper have developed over many years and been influenced by discussions with many colleagues, notably John Andreae, Bill Clancey, Joe Goguen, Bill Gray, Richard Gregory, George Klir, Gordon Pask, Francisco Varela, Alan Watson and Ian Witten. I am grateful to the referees for perceptive critiques that greatly improved the paper. I am grateful to Mildred Shaw for very detailed comments on the paper, and for access to her own research which has heavily influenced the content.

REFERENCES

Aarts, E., & Aarts, K. J. (1989). *Simulated annealing and Boltzmann machines.* Chichester, UK: Wiley.

Ashby, W. R. (1952). *Design for a brain.* London, UK: Chapman & Hall.

Baars, B. J. (1986). *The cognitive revolution in psychology.* New York: The Guilford Press.

Barrow, J. D., & Tipler, F. J. (1986). *The anthropic cosmological principle.* Oxford: Clarendon Press.

Basar, E. (Ed.). (1990). *Chaos in brain function.* Berlin: Springer.

Bickerton, D. (1990). *Language and species.* Chicago, IL: University of Chicago Press.

Blumberg, H. (Ed.). (1983). *The legitimacy of the modern age.* Cambridge, MA: MIT Press.

Blume, S. S. (1974). *Toward a political sociology of science.* New York: Free Press.

Blumer, H. (1969). *Symbolic interactionism: Perspective and method.* Englewood Cliffs, NJ: Prentice-Hall.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process.* Chicago, IL: University of Chicago Press.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Cannon, W. B. (1932). *The wisdom of the body.* London: Chapman & Hall.

Castoriadis, C. (1987). *The imaginary institution of society.* Cambridge, MA: MIT Press.

Clancey, W. J. (1989). Viewing knowledge bases as qualitative models. *IEEE Expert, 4*(2), 9–23.

Conant, R. C., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal Systems Science, 1*(2), 89–97.

Costall, A., & Still, A. (Eds.). (1987). *Cognitive psychology in question.* New York: St. Martin's Press.

Crespi, F. (1989). *Social action and power*. Oxford: Basil Blackwell.
Dawkins, R. (1982). *The extended phenotype*. Oxford: Oxford University Press.
Dendrinos, D. S., & Sonis, M. (1990). *Chaos and socio-spatial dynamics*. Berlin: Springer.
Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.
Fodor, J. (1986). The modularity of mind. In Z. W. Pylyshyn & W. Demopoulos (Eds.), *Meaning and cognitive structure* (pp. 3–18). Norwood, NJ: Ablex.
Gaines, B. R. (1969). Linear and nonlinear models of the human controller. *International Journal of Man-Machine Studies, 1*(4), 333–360.
Gaines, B. R. (1976). Behaviour/structure transformations under uncertainty. *International Journal Man-Machine Studies, 8*(3), 337–365.
Gaines, B. R. (1977). System identification, approximation and complexity. *International Journal of General Systems, 2*(3), 241–258.
Gaines, B. R. (1987). Positive feedback processes underlying functional differentiation. In M. Caudhill & C. Butler (Eds.), *Proceedings of IEEE First International Conference on Neural Networks* Vol. 2, pp. 387–394. New York, NY.
Gaines, B. R. (1988). Positive feedback processes underlying the formation of expertise. *IEEE Transactions on Systems, Man & Cybernetics, SMC-18*(6), 1016–1020.
Gaines, B. R. (1989). Social and cognitive processes in knowledge acquisition. *Knowledge Acquisition, 1*(1), 251–280.
Gaines, B. R. (1991). Between neuron, culture and logic: Explicating the cognitive nexus. *ICO: Intelligence Artificielle et Sciences Cognitives au Québec, 3*(2), 47–61.
Gaines, B. R., & Shaw, M. L. G. (1984). Hierarchies of distinctions as generators of system theories. In A. W. Smith (Ed.), *Proceedings of the society for general systems research international conference* (pp. 559–566). Louisville, KY: Society for General Systems Research.
Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
Geuss, R. (1981). *The idea of a critical theory: Habermas and the Frankfurt School*. Cambridge, UK: Cambridge University Press.
Gibson, J. J. (1979). *The ecological approach to perception*. Boston, MA: Houghton Mifflin.
Goodman, N. (1973). *Fact, fiction and forecast*. Indianapolis, IN: Bobbs-Merrill.
Gould, S. J. (1989). *Wonderful life: The Burgess shale and the nature of history*. New York: Norton.
Gray, W. (1979). Understanding creative thought processes: An early formulation of the emotional-cognitive structure theory. *Man-Environment Systems, 9*, 3–14.
Hagstrom, W. O. (1965). *The scientific community*. New York: Basic Books.
Hall, E. T. (1959). *The silent language*. New York: Doubleday.
Hayles, N. K. (Ed.). (1991). *Chaos and order: Complex dynamics in literature and science*. Chicago: University of Chicago Press.
Hirst, W. (1988). *The making of cognitive science*. Cambridge, UK: Cambridge University Press.
Homans, G. C. (1974). *Social behavior: Its elementary forms*. New York: Harcourt Brace Jovanovich.
Homans, G. C. (1987). Behaviourism and after. In A. Giddens & J. Turner (Eds.), *Social theory today* (pp. 58–81). Stanford, CA: Stanford University Press.
Hume, D. (1739). *A treatise of human nature*. Oxford, UK: Clarendon Press.
Jantsch, E. (1980). *The self-organizing universe*. Oxford, UK: Pergamon Press.
Johnson-Laird, P. N., & Wason, P. C. (Eds.). (1977). *Thinking: readings in cognitive science*. Cambridge, UK: Cambridge University Press.
Kant, I. (1781). *Critique of pure reason* (J. M. D. Meiklejohn, Trans.). London, UK: George Bell.
Kelly, G. A. (1955). *The psychology of personal constructs*. New York: Norton.

Klir, G. J. (1976). Identification of generative structures in empirical data. *International Journal of General Systems, 3*, 89–104.

Klir, G. J. (1985). *Architecture of systems problem solving*. New York: Plenum Press.

Lefebvre, V. A. (1982). *Algebra of conscience*. Dordrecht, Holland: Reidel.

Levis, A. J. (1977). The formal theory of behavior. *Journal of Social Psychiatry, 23*(2).

Luhmann, N. (1979). *Trust and power*. Chichester, UK: Wiley.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behaviour*. Oxford: Pergamon Press.

Maturana, H. R. (1975). The organization of the living: A theory of the living organization. *International Journal Man-Machine Studies, 7*(3), 313–332.

Maturana, H. R. (1981). Autopoiesis. In M. Zeleny (Ed.), *Autopoiesis: A theory of living organization* (pp. 21–33). New York: North Holland.

McCoy, M. M. (1981). Positive and negative emotion: A personal construct theory interpretation. In H. Bonarius, R. Holland, & S. Rosenberg (Eds.), *Personal construct psychology: Recent advances in theory and practice* (pp. 95–104). London, UK: MacMillan.

Melges, F. T. (1982). *Time and the inner future: A temporal approach to psychiatric disorders*. New York: Wiley.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56–63.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago, IL: University of Chicago Press.

Meyering, T. C. (1989). *Historical roots of cognitive science*. Dordrecht, Holland: Kluwer.

Monod, J. (1972). *Chance and necessity*. London, UK: Collins.

Ollason, J. G. (1991). What is this stuff called fitness? *Biology and Philosophy, 6*(1), 81–92.

Pask, G. (1975). *Conversation, cognition and learning*. Amsterdam: Elsevier.

Pepper, S. C. (1942). *World hypotheses*. Los Angeles, CA: University of California Press.

Pope, S. (1984). Conceptual synthesis: Beating at the ivory gate? In W. Smith (Ed.), *Systems methodologies and isomorphies* (pp. 31–40). Berkeley, CA: Intersystems.

Popper, K. R. (1968). Epistemology without a knowing subject. In B. V. Rootselaar (Ed.), *Logic, Methodology and Philosophy of Science III* (pp. 333–373). Amsterdam: North-Holland.

Popper, K. R. (1974). Autobiography of Karl Popper. In P. A. Schilpp (Ed.), *The Philosophy of Karl Popper* (pp. 3–181). La Salle, IL: Open Court.

Prigogine, I. (1984). *Order out of chaos*. Toronto, Ont.: Bantam.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Ruelle, D. (1989). *Chaotic evolution and strange attractors*. Cambridge, UK: Cambridge University Press.

Sharif, N. (1978). An introduction to modeling. In N. Sharif & P. Adulbhan (Eds.), *Systems modeling for decision making* (pp. 1–21). Bangkok: Asian Institute of Technology (distributed by Pergamon Press).

Shaw, M. L. G. (1985). Communities of knowledge. In F. Epting & A.W. Landfield (Eds.), *Anticipating personal construct psychology* (pp. 25–35). Lincoln, NE: University of Nebraska Press.

Simmel, G. (1950). *The sociology of Georg Simmel* (Wolff, K. H., Trans.). New York: Free Press.

Skinner, B. F. (1974). *About behaviourism*. London, UK: Jonathan Cape.

Skinner, B. F. (1989). *Recent issues in the analysis of behavior*. Columbus, OH: Merrill.

Stich, S. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.

Stich, S. P., & Nisbett, R. E. (1984). Expertise, justification and the psychology of inductive reasoning. In T. L. Haskell (Ed.), *The authority of experts* (pp. 226–241). Bloomington, IN: Indiana University Press.

Touraine, A. (1965). *La sociologie de l'action*. Paris: Éditions du Seiul.

Tucker, D. M., & Williamson, P. A. (1984). Asymmetric neural control systems in human self-regulation. *Psychological Review, 91*(2), 185–215.

Vygotsky, L. S. (1934). *Thought and language*. Cambridge, MA: MIT Press.

Wallace, R. A., & Wolf, A. (1991). *Contemporary sociological theory*. Englewood Cliffs, NJ: Prentice-Hall.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158–177.

Wolff, K. H. (1976). *Surrender and catch: Experience and enquiry today*. Holland: Reidel.

Address all correspondence to: Brian R. Gaines, Knowledge Science Institute, University of Calgary, Calgary, Alberta, Canada T2N 1N4; email gaines@cpsc.ucalgary.ca.